BCSE334L – Predictive Analytics

Case Study

Consider data/dataset in the domain as stated in the table (given below) and perform the following on the data/dataset in any programming language of your choice. Use a benchmarked dataset/ generate a dataset from a standard web domain/synthesize your own dataset.

i) Apply appropriate pre-processing techniques that is applicable for your dataset

ii) Apply any two Classification/Regression algorithms.

ii) Apply any one Clustering algorithm.

NOTE:

Your question number = (Last_4_characters_of_your_register_number mod 24) + 1

The file to be uploaded should contain the following:

➔ Register Number, Name, Slot, Course Name

➔ Question Text

➔ URL of the data/dataset

➔ Introduction

➔ Problem Statement

➔ Methodology / Approach

➔ Results / Snapshot of output

➔ Conclusion

➔ Source Code

Question No.	Domain/Area of the dataset
1	Computer Networks
2	Social Media/Networks
3	Weather
4	Stack Overflow
5	Agriculture
6	Web Security

Question No.	Domain/Area of the dataset
7	Trade Analysis
8	GitHub
9	Poverty
10	Crops
11	Sports
12	Human Resources
13	LinkedIn
14	Water
15	E-Shopping
16	Global Migration
17	Facebook
18	Health Care
19	Banking
20	Defect Detection
21	Soil Types
22	Programming Languages
23	Education
24	NGO